

Rochester Institute of Technology RIT Scholar Works

Articles

7-1993

The Next Generation of Human-Computer Interfaces: A Psychological Perspective

Morton Isaacs

Rochester Institute of Technology

Follow this and additional works at: <http://scholarworks.rit.edu/article>

Recommended Citation

Isaacs, M. (1993). The next generation of human-computer interfaces: A psychological perspective. Paper presented at the International Workshop on the Next Generation of Information Technologies and Systems, Haifa, Israel, July.

This Article is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Articles by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

THE NEXT GENERATION OF HUMAN-COMPUTER INTERFACES: A PSYCHOLOGICAL PERSPECTIVE

Paper presented as a panel discussant at The International Workshop
on the Next Generation of Information Technologies and Systems.
Haifa, Israel

July 1993

Morton Isaacs, Ph.D., Professor of Psychology, Dept of Behavioral
Science Rochester Institute of Technology, Rochester, New York 14623

- - - - -

Before we can extrapolate to the next generation of human-computer
interfaces, a relevant question is "How many generations of human-
computer interfaces have there been so far, and what were they?"

I believe the question could be answered this way from a psychological
perspective:

Generation 1: Humans threw switches manually to alter the flow
of information (original ENIAC computer).

Generation 2: Humans wrote machine language programs which
set the switches electronically.

Generation 3: Humans wrote assembly language programs
which allowed more human-like words that the machine translated into
machine language.

Generation 4: Humans used "higher level languages" like BASIC
and ALGOL to write programs which could be translated by a computer
into groups of machine language commands. MS-DOS as an operating
system would fit in here.

Generation 5: Pointing replaced or supplemented typing as
input. MS-DOS Windows and the Macintosh exemplify this generation.

What trends in human/computer interface do the above indicate?

First, each shifted more physical work onto the computer,
lessening the chance for human fatigue leading to errors.

Second, there is a continuous simplification in the amount and quality of human memory necessary to have the computer perform desired actions.

Third, there is an attempt to lessen or eliminate the need for skillfulness in the input device as much as possible.

- - - - -

Let's examine why these changes have occurred.

Work: The original purpose of all machines was to lessen the physical labor of humans; that has not changed. Few people would look forward to having to write the same information over and over; the invention of carbon paper, then copying machines, now e-mail permitting unlimited copies of information with little work on the part of the copier illustrates this progress.

Memory: Human memory consists of three separate systems--sensory storage, short-term memory, and long-term memory. Only in sensory storage is memory kept as an exact replica of the input and that for only 250 msec. In short-term memory, the input is modified by being chunked, organized, and labeled, influencing what can be retrieved; these memory traces last only about 25 secs. and then disappear for good. Information is coded in long-term memory by being categorized: certain stimulus aspects are sharpened while others are leveled. Previous memories can distort or totally interfere with current memories (proactive interference), while current memories can also interfere with previously retained information (retroactive interference).

It is in terms of exact retention of information that computers drastically outperform humans; a computer can retain millions of unrelated bits of information for years without noticeable deterioration. The less new information required to use computers, therefore, the fewer the errors that will be made by the human user. Humans forced to learn new or specific terms for common actions to interact with computers often react with frustration, anger, and input errors.

Skillfulness: Skill is often in short supply and therefore valuable. A fast, accurate typist is hard to find. Many people do not know how to touch-type and therefore are very slow at typing input, even those who do touch-type often make mistakes while typing, and there are

few among us who haven't felt fingers cramp or wrists ache after a long day at present-day computers. Typing is an inefficient method of communicating information -- like finger-spelling for the deaf versus signing, it should only be used if total precision is desired and then for short periods of time. In typing each individual finger motion must be repeated again and again although finger muscles are not mechanically designed to do this, leading to fatigue and eventually R.M.I. (Repetitive Motion Injury). Eyes are not meant to be focused for long periods of time on close objects as they must be when inputting information through typing and headaches and eyestrain often result. The body is not meant to be continuously in a sitting position for long periods of time, resulting in back strain and other physical problems.

The creation of the "mouse" as an input device helped many of these things. Rather than being forced to memorize and reproduce different combinations of commands, icons are displayed that can be pointed to and which serve as a replacement for a series of typed commands. This lessens the load on all three of the factors of work, memory, and skillfulness. The mouse however has drawbacks of its own: it is limited to choosing what is displayed on a crowded screen; it still requires repetitive finger motion and offers the chance for errors although fewer than in typing; and it is slow.

The Next Generation...

It is clear to me that input devices must move toward methods that require even less work, are less memory-intensive, and demand even less skill than typing. The spoken language is so natural a communicative medium that I believe vocal input is the method which will predominate within ten years in all computer systems. The reasons why are these:

ADVANTAGES OF VOCAL INPUT

(1) All normal children acquire a fairly extensive language capability in their native language by the age of five. According to Chomsky, a Language Acquisition Device is part of our mental equipment from birth. Even if this is not true, it is clear that children acquire language quickly and painlessly in normal situations. Vocal capabilities do not require anywhere near as much work on the child's part as learning to type, nor does it consume memory, nor is it particularly skillful. The potential input market for computer use will magnify many-fold when computers can understand verbal commands.

(2) We make many fewer errors when talking than when typing. Although I may type the word "copy" as "coyp" I would never say it that way. If I do make a mistake while talking, I am most often aware of the mistake and can correct it immediately which is not as true of typing.

(3) Vocal output can long outlast finger movements without stress or difficulty. I have been talked to by a person for hours, often without the person permitting interruption or pause in the verbal production;; few would wish to type by the hour in the same condition.

DISADVANTAGES OF VOCAL INPUT

Are there any disadvantages to using vocal input as the major input device? Yes, there are, but all are answerable:

(1) Computers have difficulty in recognizing words and often produce nonsense instead of what is expected.

Word recognition programs have recently taken giant leaps forward to such an extent that they are now commercially available. There is no question that those presently available do make errors; there seems to be no practical reason, however, why improvements will not occur to make them much more reliable in the near future.

A subcategory of word recognition, language translation, is a less sanguine possibility as anyone who has had contact with these A.I. programs would admit. However, much progress has occurred as a consequence of increasing memory and reduced cost per chip. As shown by the news article from the NY Times, language translation programs are now being commercially produced that are just on the edge of practicality -- and may be expected to become useful in the next five years.

(2) Noise can interfere with the reception of vocal commands.

Subvocal microphones are currently available which can pick up sound directly from the larynx, thereby excluding extraneous noise from the input.

(3) Vocal input is limited: with typing you can underline, italicize, bold print, and use other printed techniques which can't be indicated to the computer verbally.

Semiotics has shown that much of our understanding of the meaning of speech reflects our reception of paralinguistic and kinesic cues of which we are not, or only marginally, aware. We increase the volume of sound on a word to indicate emphasis, pause to put quotes around a word, etc. At the present time, computers are not receptive to these input variables but there is no theoretical reason why they could not be recognized and utilized in a program.

(4) There are situations where vocal input would be difficult or impossible. For example, dictation is a difficult skill to master as compared to typing.

Certainly, there are situations and various people for whom vocal input will not be possible or desired. There will always be a place for typing or a mouse in certain computer situations: graphics design, for example, or special effects with fonts. Those who are mute either temporarily or permanently, those who feel extremely uncomfortable talking rather than writing, and those whose language is so non-standard that the computer might not be able to understand them all might desire or wish to input in writing/typing/mousing rather than by vocal command. At present one can input into a computer by importing data from some other source; I see no reason why these methods will not still be available for those who desire or need them for some reason.

CONCLUSION

I believe that, in general, vocal input will be the usual input method in the next generation for reasons of ease of use, ready availability of cue words in natural language, and its lack of skill needed for successful computer interaction. It now behooves the software and hardware technicians to work to make it as transparent, accurate, and user-friendly as possible.